Structural determination of lipid antigens captured at the CD1d-T-cell receptor interface

Brennan, Patrick J; Cheng, Tan-Yun; Pellicci, Daniel G; Watts, Gerald F M; Veerapen, Natacha; Young, David C; Rossjohn, Jamie; Besra, Gurdyal S; Godfrey, Dale I; Brenner, Michael B; Moody, D Branch

DOI:
10.1073/pnas.1705882114

Classification: Biological sciences

Title
Structural determination of lipid antigens captured at the CD1d-T cell receptor interface

Short Title
Capture of lipid antigens

Authors
Patrick J. Brennan[a,1,3], Tan-Yun Cheng[a,1], Daniel G. Pellicci[b,c,1], Gerald F.M. Watts[a], Natacha Veerapen[d], David C. Young[a], Jamie Rossjohn[e,f,g], Gurdyal S. Besra[d], Dale I. Godfrey[b,c,2], Michael B. Brenner[a,2,3], D. Branch Moody[a,2,3]

[a]Department of Medicine, Division of Rheumatology, Immunology and Allergy, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA.
[b]Department of Microbiology and Immunology, Peter Doherty Institute for Infection and Immunity, University of Melbourne, Parkville, Australia.
[c]ARC Centre of Excellence in Advanced Molecular Imaging, University of Melbourne, Parkville, Australia.
[d]School of Biosciences, University of Birmingham, Birmingham B15 2TT, UK.
[e] Infection and Immunity Program and Department of Biochemistry and Molecular Biology, Biomedicine Discovery Institute, Monash University, Clayton, Victoria 3800, Australia
[f]Institute of Infection and Immunity, Cardiff University School of Medicine, Heath Park, Cardiff CF14 4XN, UK
[g]ARC Centre of Excellence in Advanced Molecular Imaging, Monash University, Clayton, Australia.

[1]These authors contributed equally to this work.
[2]These authors jointly directed this work.
[3]To whom correspondence should be addressed.

Corresponding authors
Michael B. Brenner
Brigham and Women's Hospital
BTM, 6th floor, room 6002U
60 Fenwood Road
Tel: 617-525-1000
Email: mbrenner@research.bwh.harvard.edu

D. Branch Moody
Brigham and Women's Hospital
BTM, 6th floor, room 6002K
60 Fenwood Road
Tel: 617-525-1037
Email: bmoody@bwh.harvard.edu

Patrick J. Brennan
Brigham and Women's Hospital
BTM, 5th floor, room 5002S
60 Fenwood Road
Tel: 617-525-1016
Email: pbrennan3@bwh.harvard.edu

Keywords

NKT
iNKT
Lipid antigen
CD1d
Glycolipid

Abstract

Glycolipid antigens recognized by αβ T cell receptors (TCRs) drive the activation of invariant natural killer T (iNKT) cells, a specialized subset of innate T lymphocytes. Glycolipids with α-linked anomeric carbohydrates have been identified as potent microbial lipid antigens for iNKT cells, and their unusual α-anomeric linkage has been thought to define a 'foreign' lipid antigen motif. Yet mammals use endogenous lipids to select iNKT cells, and there is compelling evidence for iNKT cell responses in various types of sterile inflammation. The nature of endogenous or environmental lipid antigens encountered by iNKT cells is not well defined. Here, we sought to identify lipid antigens in cow's milk, a prominent part of the human diet. We developed a method to directly capture lipid antigens within CD1d-lipid-TCR complexes, while excluding CD1d-bound to non-antigenic lipids, followed by direct biochemical analysis of the lipid antigens trapped at the TCR-CD1d interface. The specific antigens captured by this 'TCR trap' method were identified as α-linked monohexosylceramides by mass spectrometry fragmentation patterns that distinguished α- from β-anomeric monohexosylceramides. These data provide direct biochemical evidence for α-linked lipid antigens from a common dietary source.

Significance Statement

Invariant natural killer T (iNKT) cells are activated quickly and play a key role in the control of many microbial infections via their ability to rapidly secrete cytokines and chemokines that enhance many immune responses. Microbial glycolipid antigens that activate iNKT cells have been identified as α-anomerically linked glycolipids. However, the nature of the endogenous lipid antigens that are important contributors to the biology of iNKT cells has been unclear. In this study, antigenic lipids from cow's milk were isolated using a T cell receptor trap method, and their stereochemical structures were determined to be hexosylceramides with α-linked hexose headgroups, thus identifying dietary lipid antigens for iNKT cells.

\body

Introduction

Adaptive mammalian T lymphocytes are well known for their use of combinatorial diversity to generate a broad repertoire of antigen receptors, a process that evolved to recognize newly-encountered peptide epitopes from

pathogens. However, certain subsets of human T cells, such as invariant natural killer T (iNKT) cells, mucosal-associated invariant T (MAIT) cells, and germline-encoded mycolyl-reactive (GEM) T cells reliably generate T cell receptors (TCRs) that are conserved among genetically diverse humans, and in some cases, across species (1). Recent data indicate that T cells expressing such invariant TCRs comprise a significant fraction of all human T cells, estimated in the range of 10 to 20 percent (2). Further, these conserved T cell subsets are not restricted by major histocompatibility (MHC)-encoded antigen presenting molecules, but instead can recognize different chemical classes of antigens presented by non-polymorphic antigen presenting molecules including CD1, MR1 or HLA-E (3). One of these conserved T cell populations, iNKT cells, recognizes specific glycolipids presented by CD1d molecules. Following activation, iNKT cells respond rapidly with vigorous cytokine production that can profoundly shape immune responses (4).

Although a glycolipid antigen for iNKT cells was identified from a marine sponge more than 20 years ago as part of a chemical library screen (5), the origins and chemical identities of physiologically-relevant antigens that drive selection and activation of iNKT cells remain controversial. Natural antigenic glycolipids have been detected from various foreign sources, including non-pathogenic α-proteobacteria (6-8), spirochetes (9), and other microbes such as *Streptococcus pneumoniae* (10), and *Bacteroides fragilis* (11, 12). These known microbial lipid antigens for iNKT cells share a key biochemical feature, namely an α-anomeric linkage of one hexose sugar to the lipid backbone. Mammals abundantly produce β-linked hexosylceramides, glucosylceramide and galactosylceramide, but were not previously thought to produce α-linked hexosylceramides. Therefore, a major principle of antigen recognition by iNKT cells has been that self-glycolipids are β-linked and not strongly recognized, while pathogens produce lipids that are recognized as foreign based on their α-anomeric linkage (5). However, pathogen-derived lipids presumably do not explain the TCR- and CD1d-dependent activation of iNKT cells during positive selection in the thymus, or in settings such as viral infection, cancer, or sterile inflammation (4). Endogenous lipids have long been proposed to activate iNKT cells in these settings.

It is possible that the longstanding view that mammalian monohexosylceramides are always β-linked is incorrect. Kain and colleagues showed that antibodies raised against α-linked glycosphingolipids could stain mammalian tissues (13). We recently fractionated mammalian milk-derived lipids, finding that antigenic substances were enriched in the hexosylceramide-containing fraction, and that antigenicity was not abrogated by β-specific glycosidases or chromatographic depletion of β-glucosylceramides (14, 15). Both studies point to the possibility that α-linked lipids are produced in mammalian tissues. However, if they exist, they are likely scarce compounds and they have not been isolated in sufficient yield to confirm α-anomeric structures through direct biochemical analysis. Given the importance of the 'α-linked self-lipid hypothesis' to iNKT cell biology, we sought to identify mammal-derived antigenic lipids.

Here, we used soluble iNKT TCR and soluble CD1d to capture the subset of lipids from cow's milk that form stable TCR-lipid-CD1d complex through direct analysis of TCR-lipid-CD1d complexes. Combining this TCR 'trapping' of CD1d-lipid antigen complexes with nanoelectrospray ionization (ESI) mass spectrometry (MS) and liquid chromatography-mass spectrometry (LC-MS), we could detect monohexosylceramides from cow's milk as lipid antigens. Using MS-based methods, we determined that the captured lipid antigens were α-linked monohexosylceramides. This study identifies the biochemical structures of natural cow's milk-derived lipid antigens as α-anomeric hexosylceramides, supporting the hypothesis that α-linked lipids are iNKT cell antigens even in the absence of pathogens.

Results

Pre-purification of milk lipid antigens for iNKT cells. We previously showed that the glucosylceramide-containing fraction of cow's milk contains antigenic lipids for iNKT cells (14, 15). To further purify and identify the milk-derived lipid antigens, we digested a milk hexosylceramide lipid fraction with β-glucosidase to destroy the β-linked glucosylceramides. Thin layer chromatography (TLC) showed that this treatment markedly

reduced the total mass of glycolipids co-migrating with a glucosylceramide standard, yet did not reduce the potency of iNKT cell activation (Fig. 1A). This result could be explained if the starting material was predominantly composed of non-antigenic β-linked glucosylceramides, and antigenic α-linked hexosylceramides were present in milk that survived the enzyme treatment. To address this possibility, we used a TLC separation method that resolved monohexosylceramide α- and β- anomers (15), yielding two bands. Only the lower band, which co-migrated with an α-glucosylceramide standard, stimulated an iNKT cell response (Fig. 1B), suggesting that α-linked hexosylceramides could be present. However, this interpretation is contrary to the prevailing view that mammalian products contain only β-linked hexosylceramides. The lower band that activated iNKT cells (Fig. 1A,B) which we refer to as the "activity-enriched" fraction, represented a very small fraction of total milk hexosylceramides, which provided an obstacle to formal biochemical identification of the antigen. Thus, we considered new approaches to antigen discovery.

Milk lipids mediate CD1d-TCR binding. Nearly all prior approaches to antigen discovery, including that described above (Fig. 1A), have relied on a correlation between two assays that separately measure T cell activation and biochemical content. The final antigen identification relies on the premise that the most readily detectable molecules are also the T cell stimulants. By definition, antigenic lipids ligate TCRs to antigen-presenting molecules with higher affinity than non-antigenic lipids, so we reasoned that a recombinant iNKT TCR with relatively high affinity for CD1d-α-galactosylceramide complexes, known as 2C12 (16), might selectively trap those CD1d complexes carrying milk-derived antigens, while leaving behind CD1d molecules that were bound to non-antigenic lipids. Such a TCR trap approach was recently used to identify CD1a autoantigens (17). We hypothesized that this approach might be feasible for CD1d and milk-derived antigens, since partially-enriched milk fractions could mediate the physical binding of CD1d tetramers to cells expressing an iNKT cell Vα14Jα18 Vβ8.2Jβ2.1 TCR, but not to cells expressing a sulfatide-reactive diverse NKT cell Vα1Jα26 Vβ16Jβ2.1 TCR (18, 19). Staining of iNKT TCR-expressing cells with CD1d tetramers loaded with activity-enriched milk lipids was stronger than staining with unloaded CD1d tetramers, but weaker than tetramers generated with known potent lipid antigens α-galactosylceramide and α-glucosylceramide (Fig. 1C). Thus, cow's milk contained compounds capable of mediating CD1d-TCR binding.

TCR Trap Assay. To develop a cell-free assay to detect lipid antigens, we tested the TCR trap system by loading recombinant soluble CD1d with synthetic α-glucosylceramide, and then mixing this antigen-loaded CD1d with recombinant soluble iNKT cell TCR monomers. The subset of CD1d proteins carrying α-glucosylceramide was expected to bind TCR heterodimers leading to the formation of a higher molecular weight ternary complex, which could be purified by size exclusion gel chromatography. A mixture of α-glucosylceramide-treated CD1d and TCRs showed a large early fraction that eluted prior to separate CD1d and TCR monomers (Fig. 2A), consistent with a larger protein complex. After eluting lipids from this fraction using a mixture of chloroform and methanol, positive mode nanoelectrospray MS identified a prominent ion at *m/z* 832.7, corresponding to α-glucosylceramide ([M+Na], M=$C_{48}H_{91}NO_8$), as well as its isotopes. The lack of other ions present above background levels indicated that lipids from culture media or those potentially carried by the outer surface of CD1d or TCRs did not substantially confound the analysis (Fig. 2B). Thus, α-glucosylceramide performed its expected role in ligating CD1d to TCR, and direct analysis of CD1d-lipid-TCR complexes was sufficient to identify the ligating molecule without interference by other substances adherent to the lipid-protein complex.

Next, to identify unknown lipids, we used this TCR trap assay with CD1d molecules that had been loaded with activity-enriched cow's milk monohexosylceramides (Fig. 1A,B). A small but detectable peak was observed in the early size exclusion fractions and this showed no overlap with CD1d or TCR monomer fractions (Fig. 2C). Coomassie staining of the early fraction revealed three proteins matching the apparent molecular weights of CD1d, β2m, and TCR with band similar intensity, consistent with 1:1:1 stoichiometry (Fig. 2D). The peak height of the protein trace of the early eluting fraction made with milk lipids (Fig. 2C) was substantially lower than that of the early fraction made with CD1d carrying synthetic α-glucosylceramide (Fig. 2A), suggesting that TCRs bound to a small fraction of the total pool of CD1d-lipid complexes present. However, since this peak was observed at the same elution time as the synthetic α-glucosylceramide standard-loaded CD1d, it suggested that the milk lipid-derived ternary complex carried a high-affinity antigen.

Nanoelectrospray analysis of all eluted milk lipids in the late fraction containing CD1d and TCR monomers found numerous ions, and the mass intervals among the most abundant ions were diverse, suggesting the presence of structurally diverse lipids from many classes (Fig. 3A, upper). In contrast, the early fraction, which contained CD1d-TCR complexes, showed fewer eluted ions (Fig. 3A, lower). The lower peak height from the size exclusion tracing and the less complex MS profiles suggested that the iNKT cell TCR selected small subset of CD1d-milk lipid complexes from among all CD1d-lipid complexes present in the starting material. In fact, nearly all nanoelectrospray MS signals derived from the early-eluting CD1d-TCR complexes could be accounted for by a single alkane series comprised of lipids of *m/z* 710.7, apparent chain-length analogs (*m/z* 724.7, 738.7, 752.6, 794.7, 808.8, 822.8, 836.8), and the isotopes of these lipids.

TCR ligands identified using nanoelectrospray MS. Next we identified the lipid ligands trapped-in or excluded-from CD1d-TCR complexes using shotgun nanoelectrospray with multi-stage ion trap MS. This method provides high sensitivity, broad detection, and structural information from collision-induced dissociation (CID), but offers relatively low mass accuracy. Considering the lipids eluted from the late fraction containing CD1d and TCR monomers, CID-MS of selected ions at *m/z* 738.4, 753.5, and 808.5 identified phosphatidylethanolamine, sphingomyelin, and phosphatidylcholine (Fig. 3B-D). In addition, fragments from the ion at *m/z* 722.6 corresponded a hexosylceramide whose two fatty acid chains had a combined lipid tail length of C34 with one unsaturation (Fig. 3E). This ‘type 1’ monohexosylceramide collision pattern was characterized by a dehydration (loss of 18 u, *m/z* 704.5), loss of a hexosyl unit (162 u, *m/z* 560.5), and loss of both fragments (loss of 180 u, *m/z* 542.4). In contrast, ions from early fraction containing CD1d-TCR complexes did not show detectable signals for phosphatidylethanolamine, sphingomyelin, or phosphatidylcholine. Instead, complexes in the early fraction released matched expected masses of a ‘type 2’ hexosylceramide, which unlike type 1 patterns, lacked a dehydration ion but did show a loss of 162 u. In fact, the abundant ions in the early fraction were all accounted for by single alkane series corresponding to an monohexosylceramide with an overall lipid length ranging between C33 (*m/z* 710.7) and C42 (*m/z* 836.8). Separate analysis of six members of this alkane series demonstrated highly reproducible ‘type 2’ collision pattern (Fig. 3F).

The detailed profiles of these released lipids suggested several general conclusions. First, monomeric CD1d proteins were bound to phospholipids, sphingomyelin, type 1 hexosylceramide, as well as other lipids that were not specifically identified (Fig. 3A). Second, TCR-CD1d complex formation largely or completely excluded CD1d bound to phosphatidylcholine (*m/z* 808.6 with the *m/z* 749.3 product), sphingomyelin (*m/z* 753.6 with *m/z* 694.4 product), phosphatidylethanolamine (*m/z* 738.5 with *m/z* 695.3 product), and the type 1 hexosylceramide with a monounsaturated lipid tail (*m/z* 722.6 with *m/z* 542.4 fragment). Third, lipids eluted from the TCR-CD1d complexes contained type 2 monohexosylceramides ranging from C33-42 in chain length (Fig. 3A). Further, the characteristic patterns of loss in CID analyses by type 1 (18, 162, and 180 u) and type 2 (162 u) hexosylceramides implicated a differing conformation of these ceramide-based lipids. This interpretation is consistent with the TCR discriminating structural classes of lipids including glycoslyceramides, and possibly α- versus β-linked hexosylceramides.

The molar concentrations of hexosylceramides eluted from the early and late complexes were determined using HPLC-MS in comparison to standards, and this material tested for the ability to elicit cytokine production from an iNKT cell hybridoma, DN32. The late fraction did not have detectable activity, while the molar activity observed in the early fraction was similar to that of a synthetic antigen for iNKT cells known as α-galactosylceramide, KRN7000 (Fig. 3G). Thus, our approach was successful as a TCR-based affinity trap for selecting antigenic glycolipids that ligate TCR to CD1d out of a complex mixture of natural lipids. The equipotency of type 2 hexosylceramides and α-galactosylceramide was consistent with the type 2 hexosylceramides being an α-linked lipid. Although only trace amounts of the type 2 monohexosylceramide could be recovered from ternary complexes of CD1d-lipid-TCR, 1D-proton nuclear magnetic resonance spectroscopy (NMR) carried out on a 900 MHz instrument identified a doublet with a chemical shit of 4.85 ppm, consistent with the anomeric proton resonance of α-monohexosylceramide (Fig. S1) (15).

Identification of ligands at high mass resolution. To further biochemically characterize the lipids that promoted the formation of TCR-lipid-CD1d complexes, we separately analyzed eluents from milk lipid-treated

protein monomers present in the late fraction, untreated CD1d, and untreated TCR (Fig. 4). Here, we took advantage of the higher mass accuracy of a quadrupole time-of-flight (Q-ToF) mass detector coupled with chromatographic separation in normal-phase HPLC-MS to confirm the identity of the key ions. Using this method, the major ions previously identified by nanoelectrospray MS were confirmed and assigned based on high mass accuracy and HPLC retention time, compared with standards as phosphatidylcholine (38.1 min, *m/z* 786.601), phosphatidylethanolamine (20.5 min, *m/z* 716.523), type 1 hexosylceramide (4.1 min, m/z 700.578), and type 2 hexosylceramide (4.1 min, m/z 702.590, 716.604 and 786.682) (Figure 4). No substantial false positive signals were derived from lipids released from the TCR, indicating that the lipids detected were carried by CD1d. As seen with the nanospray method (Fig. 3), CD1d alone and mixed CD1d and TCR monomers bound phosphatidylcholine, phosphatidylethanolamine, and type 1 monohexosylceramides (Fig. 4 and S2), but these were excluded from CD1d-TCR complexes. Conversely, type 2 hexosylceramides eluted with low or undetectable signals from milk lipid treated monomers, but were highly enriched in CD1d-TCR complexes. Thus, CD1d bound many lipids, but the iNKT TCRs selectively captured CD1d bound to type 2 hexosylceramides (Fig. S3).

Structural identification of type 1 and 2 hexosylceramides. Despite their strikingly-different patterns of binding within, or being excluded from, CD1d-TCR complexes, normal-phase HPLC-MS showed that type 1 and 2 hexosylceramides had equivalent retention time and the same underlying molecular composition, except for variations in lipid tail length and unsaturation. CID-MS experiments using the ion trap architecture, however showed different fragmentation patterns for type 1 and type 2 hexosylceramides (Fig. 3). We next asked whether anomeric variants could be distinguished by CID-MS. First, we produced synthetic standards differing only in the stereochemistry of their hexose attachment. Both glucosyl and galactosyl variants were synthesized, with dihydrosphingosine (d18:0) as their lipid backbones based on the molecular composition of type 2 hexosylceramides identified in the milk-derived early fraction. These synthetic standards had a similar retention time by normal-phase HPLC to the early-eluting milk lipids (Fig. S4). The synthetic standards were tested for their ability to activate the DN32 iNKT cell hybridoma after loading in CD1d-expressing RAW264.7

cells. For both glucosyl and galactosyl synthetic standards, activity was observed for the α-linked variants, but not for the β-linked variants (Fig. 5A).

Positive mode ion trap CID-MS of sodiated adducts of these synthetic hexosylceramides was carried out under conditions that matched those used for milk-derived compounds (Fig. 3), and over a range of collision energy settings (Fig. 5B). Collision of the α-glucosylceramide synthetic standard demonstrated almost exclusively a primary loss of 162 u, matching the pattern for type 2 milk-derived hexosylceramides. In contrast, β-glucosyl standards matched the fragmentation pattern seen for type 1 hexosylceramides from milk, showing a loss of 18 u at a similar frequency as loss of 162 u. Although both patterns involved the loss of 162 u as a hexosyl unit, analysis over a range of collision energy demonstrated that β-glucosylceramide also lost 18 u and 180 u, whereas α-glucosylceramide showed neither of these fragmentations at a substantial frequency. Parallel analyses of α- and β-galactosylceramide standards showed similar characteristic α- and β- collision patterns (Fig. S5). Matched synthetic α- and β-glucosylceramide and galactosylceramide standards based on the type 1 hexosylceramide lipid backbone structure showed similar characteristic α- and β- collision patterns, demonstrating that the anomeric linkage discrimination by this technique was not limited to dihydrosphingosine-based glycosphingolipids (Fig. S6). Ion trap CID-MS of α-galactosylceramide with a phytosphingosine base also fragmented with the pattern matching type 2 hexosylceramides. We concluded from these experiments that linear ion trap MS could sensitively distinguish α- from β-hexosylceramides under the conditions measured for a range of lipids with differing tail lengths and unsaturations.

We compared the collision patterns of the major hexosylceramide ions from cow's milk lipids. Unfractionated cow's milk glucosylceramides demonstrated a fragmentation pattern characteristic of β-hexosylceramide standards (Fig. 5C), consistent with the main lipid in the starting material being β-glucosylceramide. The activity-enriched cow's milk glucosylceramide showed a β-hexosylceramide collision pattern, but with an increased frequency of a type 2 pattern characterized by the loss of the sugar headgroup (162 u). The antigenic lipids captured in the TCR trap assay showed the fragmentation pattern typical of α-

hexosylceramides. We concluded from these experiments that the TCR trap captured type 2 α-linked hexosylceramides from cow's milk. To address the possibility that these type 2 hexosylceramides were derived from microbes present in cow's milk, we purified the monohexosylceramide fraction from calf thymus, a mammalian lipid source previously shown activate iNKT cells (15). Several ions with the *m/z* seen in cow's milk were also present in calf thymus, including ions corresponding to type 2 hexosylceramides *m/z* 724.7 and 808.8 (Fig. S7). Linear ion trap CID-MS of untreated calf thymus hexosylceramides showed a characteristic β-hexosylceramide pattern. After activity enrichment of thymus hexosylceramides by treatment with β-glucosidase and TLC fractionation as in Fig. 1, we observed a collision pattern consistent with the presence of α-hexosylceramides, based on a substantial increase in the abundance of the ion representing a primary loss of the sugar headgroup.

Discussion

Uncovering the specific antigens recognized by iNKT cells is central to understanding their roles in health and disease. Although lipid antigens for iNKT cells have been identified in the microbial world (6-10), experimental evidence has demonstrated important roles for endogenous and environmental lipid antigens. Several candidate endogenous lipid antigens have been proposed (13, 14, 20-22), though technical limitations in the ability to detect these antigens has hampered our ability to determine their contextual roles. Here, we provide direct biochemical evidence for α-hexosylceramide derived from a mammalian source, cow's milk.

There is increasing evidence that iNKT cells encounter lipid antigens at mucosal surfaces. CD1-dependent immunomodulatory lipids have been identified in gut microbes (11, 12), pollen (23), house dust (24), and in our work, cow's milk. We have previously demonstrated that cow's milk contains substantially more lipid antigen activity than human or mouse milk. Although native triglycerides are removed from infant formula during processing, bovine polar lipids, including lipid antigens for iNKT cells, are retained in the final product (15). Studies in germ-free mice suggest that the iNKT cell compartment can be shaped by CD1d-lipid signals in the enteral tract (11, 25), and thus early-life exposure to cow's milk-based infant formula might similarly shape human mucosal iNKT cells in early life. In adult mice, orally-delivered antigen activates iNKT cells in the gut

(26), suggesting that exposure to dietary lipid antigens in adult life may also have consequences. Interestingly, it is unclear whether cows have iNKT cells (27), as their CD1d does not present the long-chain α-galactosylceramide commonly used as a tool in mice and humans (28). Bovine CD1d can present short-chain α-galactosylceramides (28, 29), and it is possible that an increased presence of α-hexosylceramdies in cow's milk and tissues could lead to thymic deletion of iNKT cells, as has been observed in mice in the presence of exogenous α-galactosylceramide (30).

Although the bulk solids in milk derive overwhelmingly from mammary glands, we cannot formally rule out that colonizing microbes synthesized, or played a role in the synthesis of the α-hexosylceramides present in cow's milk. Several features of our data argue against possible contamination of milk by microbial products. We have previously observed iNKT cell-stimulating activity in the monohexosylceramide fractions of the milks, sera, and internal organs of multiple animal species (14, 15). The *N*-acyl chain usage of the α-hexosylceramides that we identified in this report mirror that of the bulk β-glucosylceramides in cow's milk, suggesting that these lipids are of mammalian origin. We also identified type 2 hexosylceramides from calf thymus with the same *m/z* as those in cow's milk. These thymic hexosylceramides, like those found in cow's milk, shifted from a β-hexosylceramide pattern toward a characteristic α-hexosylceramide collision pattern after activity enrichment. Currently, the level of α-hexosylceramide identified from cow's milk, a relatively abundant and potent natural source of lipid antigen, is near the limit of detection for the TCR trap assay system. However, we are continuing to improve the system so that lipids can be captured and identified from smaller amounts of source material.

The stereochemistry of organic molecules can have significant biological consequences. Determination of stereochemistry has largely been accomplished by NMR. However, for rare compounds such as endogenous iNKT cell lipid antigens, purification of sufficient material for complete, multi-dimensional NMR analysis is currently not possible. As shown in this report and others (31, 32), CID-MS can also be used to determine stereochemical features. We tested α- and β-linked synthetic standards with both sphingosine and dihydrosphingosine bases, and found consistent characteristic α- and β- fragmentation patterns, showing that

this detection method was not limited to certain lipid backbones. Interestingly, these collision patterns were only observed with selected positive mode adducts, including sodium and rubidium. We hypothesize that for the β-hexosylceramide structure, positive adducts with a sufficiently large valence, such as sodium and rubidium, can coordinate elements on both sides of the glycoside bond, stabilizing the glycoside bond during collision, thus reducing the propensity for cleavage of the sugar headgroup. For α-hexosylceramide, this headgroup stabilization may be sterically hindered, thus allowing for fragmentation across the glycoside bond at a lower energy.

Identification of CD1d-presented lipid antigens has largely relied on indirect evidence such as extrapolation from cells deficient in lipid synthesis genes (22, 33), or fractionation of crude lipid sources with activity tracking within these fractions (14, 20). These approaches have been useful, but do not directly demonstrate recognition by the TCR, and rely on correlations between two or more assays. Here, we used a direct method for antigen discovery, the TCR trap assay. This approach uses ternary complex formation both as the preparative method to recover lipids of interest and as the assay demonstrating the lipid's role in ligating CD1d and TCR. The TCR trap assay has also been used successfully to identify CD1a ligands (17), and could be used to identify ligands for other CD1 isoforms, or even MR1. This approach will allow tracking of rare lipids from virtually any source, including from complex mixtures, or limited starting materials, which represent common challenges when starting with mammalian tissues and cells.

Methods and materials

Initial lipid activity enrichment was performed as previously described (15). Mass spectrometry analyses were performed on a Thermo-Fisher LXQ ion trap mass spectrometer or an Agilent 6520 Accurate-Mass Q-ToF with HPLC. For details of experimental conditions and analysis, see SI Materials and Methods.

Acknowledgements. We thank C. Costello, C. Heiss, and P. Azadi for helpful discussions and J. Buter for assistance with figure production. P.J.B. was supported by the NIH (AI102945) and generous support from the Violin and Karol families, D.B.M. (AR048632, AI111224, and the Bill and Melinda Gates Foundation), M.B.B.

NIH (AI113046, AI063428), G.S.B. J. Badrick, the Royal Society, and the Medical Research Council, DIG and DGP were supported the National Health and Medical Research Council of Australia ((NHMRC) 1013667, 1113293, 1020770, 1117766 and 1054431) and the Australian Research Council (ARC; CE140100011), JR by the Australian Research Council (FL160100049, CE140100011).

References

1. Godfrey DI, Uldrich AP, McCluskey J, Rossjohn J, & Moody DB (2015) The burgeoning family of unconventional T cells. *Nat Immunol* 16(11):1114-1123.
2. Young MH & Gapin L (2011) Group 1 CD1-restricted T cells take center stage. *Eur J Immunol*.
3. Van Rhijn I, Godfrey DI, Rossjohn J, & Moody DB (2015) Lipid and small-molecule display by CD1 and MR1. *Nat Rev Immunol* 15(10):643-654.
4. Brennan PJ, Brigl M, & Brenner MB (2013) Invariant natural killer T cells: an innate activation scheme linked to diverse effector functions. *Nat Rev Immunol* 13(2):101-117.
5. Kawano T*, et al.* (1997) CD1d-restricted and TCR-mediated activation of valpha14 NKT cells by glycosylceramides. *Science* 278(5343):1626-1629.
6. Kinjo Y*, et al.* (2005) Recognition of bacterial glycosphingolipids by natural killer T cells. *Nature* 434(7032):520-525.
7. Mattner J*, et al.* (2005) Exogenous and endogenous glycolipid antigens activate NKT cells during microbial infections. *Nature* 434(7032):525-529.
8. Sriram V, Du W, Gervay-Hague J, & Brutkiewicz RR (2005) Cell wall glycosphingolipids of Sphingomonas paucimobilis are CD1d-specific ligands for NKT cells. *Eur J Immunol* 35(6):1692-1701.
9. Kinjo Y*, et al.* (2006) Natural killer T cells recognize diacylglycerol antigens from pathogenic bacteria. *Nat Immunol* 7(9):978-986.
10. Kinjo Y*, et al.* (2011) Invariant natural killer T cells recognize glycolipids from pathogenic Gram-positive bacteria. *Nat Immunol* 12(10):966-974.
11. An D*, et al.* (2014) Sphingolipids from a symbiotic microbe regulate homeostasis of host intestinal natural killer T cells. *Cell* 156(1-2):123-133.
12. Wieland Brown LC*, et al.* (2013) Production of alpha-galactosylceramide by a prominent member of the human gut microbiota. *PLoS Biol* 11(7):e1001610.
13. Kain L*, et al.* (2014) The identification of the endogenous ligands of natural killer T cells reveals the presence of mammalian alpha-linked glycosylceramides. *Immunity* 41(4):543-554.
14. Brennan PJ*, et al.* (2011) Invariant natural killer T cells recognize lipid self antigen induced by microbial danger signals. *Nat Immunol* 12(12):1202-1211.
15. Brennan PJ*, et al.* (2014) Activation of iNKT cells by a distinct constituent of the endogenous glucosylceramide fraction. *Proc Natl Acad Sci U S A* 111(37):13433-13438.
16. Burdin N*, et al.* (1998) Selective ability of mouse CD1 to present glycolipids: alpha-galactosylceramide specifically stimulates V alpha 14+ NK T lymphocytes. *J Immunol* 161(7):3271-3281.
17. Birkinshaw RW*, et al.* (2015) alphabeta T cell antigen receptor recognition of CD1a presenting self lipid ligands. *Nat Immunol* 16(3):258-266.
18. Cardell S*, et al.* (1995) CD1-restricted CD4+ T cells in major histocompatibility complex class II-deficient mice. *J Exp Med* 182(4):993-1004.
19. Patel O*, et al.* (2012) Recognition of CD1d-sulfatide mediated by a type II natural killer T cell antigen receptor. *Nat Immunol* 13(9):857-863.
20. Facciotti F*, et al.* (2012) Peroxisome-derived lipids are self antigens that stimulate invariant natural killer T cells in the thymus. *Nat Immunol* 13(5):474-480.
21. Fox LM*, et al.* (2009) Recognition of lyso-phospholipids by human natural killer T lymphocytes. *PLoS Biol* 7(10):e1000228.

22. Zhou D, *et al.* (2004) Lysosomal glycosphingolipid recognition by NKT cells. *Science* 306(5702):1786-1789.
23. Agea E, *et al.* (2005) Human CD1-restricted T cell recognition of lipids from pollens. *J Exp Med* 202(2):295-308.
24. Wingender G, *et al.* (2011) Invariant NKT cells are required for airway inflammation induced by environmental antigens. *J Exp Med* 208(6):1151-1162.
25. Olszak T, *et al.* (2012) Microbial exposure during early life has persistent effects on natural killer T cell function. *Science* 336(6080):489-493.
26. Lee YJ, *et al.* (2015) Tissue-Specific Distribution of iNKT Cells Impacts Their Cytokine Response. *Immunity* 43(3):566-578.
27. Van Rhijn I, *et al.* (2006) The bovine CD1 family contains group 1 CD1 proteins, but no functional CD1d. *J Immunol* 176(8):4888-4893.
28. Wang J, *et al.* (2012) Crystal structures of bovine CD1d reveal altered alphaGalCer presentation and a restricted A' pocket unable to bind long-chain glycolipids. *PLoS One* 7(10):e47989.
29. Nguyen TK, *et al.* (2013) The bovine CD1D gene has an unusual gene structure and is expressed but cannot present alpha-galactosylceramide with a C26 fatty acid. *Int Immunol* 25(2):91-98.
30. Pellicci DG, *et al.* (2003) Intrathymic NKT cell development is blocked by the presence of alpha-galactosylceramide. *Eur J Immunol* 33(7):1816-1823.
31. Kanie O, *et al.* (2009) Analysis of behavior of sodiated sugar hemiacetals under low-energy collision-induced dissociation conditions and application to investigating mutarotation and mechanism of a glycosidase. *Proc Jpn Acad Ser B Phys Biol Sci* 85(6):204-215.
32. Moody DB, *et al.* (2000) CD1c-mediated T-cell recognition of isoprenoid glycolipids in Mycobacterium tuberculosis infection. *Nature* 404(6780):884-888.
33. Darmoise A, *et al.* (2010) Lysosomal alpha-galactosidase controls the generation of self lipid antigens for natural killer T cells. *Immunity* 33(2):216-228.
34. Jervis PJ, *et al.* (2010) Synthesis and biological activity of alpha-glucosyl C24:0 and C20:2 ceramides. *Bioorg Med Chem Lett* 20(12):3475-3478.
35. Uldrich AP, *et al.* (2013) CD1d-lipid antigen recognition by the gammadelta TCR. *Nat Immunol* 14(11):1137-1145.
36. Matsuda JL, *et al.* (2000) Tracking the response of natural killer T cells to a glycolipid antigen using CD1d tetramers. *J Exp Med* 192(5):741-754.
37. Pellicci DG, *et al.* (2009) Differential recognition of CD1d-alpha-galactosyl ceramide by the V beta 8.2 and V beta 7 semi-invariant NKT T cell receptors. *Immunity* 31(1):47-59.
38. Lantz O & Bendelac A (1994) An invariant T cell receptor alpha chain is used by a unique subset of major histocompatibility complex class I-specific CD4+ and CD4-8- T cells in mice and humans. *J Exp Med* 180(3):1097-1106.
39. Muindi K, *et al.* (2010) Activation state and intracellular trafficking contribute to the repertoire of endogenous glycosphingolipids presented by CD1d [corrected]. *Proc Natl Acad Sci U S A* 107(7):3052-3057.
40. Bligh EG & Dyer WJ (1959) A rapid method of total lipid extraction and purification. *Canadian journal of biochemistry and physiology* 37(8):911-917.
41. Layre E, *et al.* (2011) A comparative lipidomics platform for chemotaxonomic analysis of Mycobacterium tuberculosis. *Chem Biol* 18(12):1537-1549.

Figure legends

Fig.1. iNKT cell activation by milk hexosylceramides. (A) A cow's milk glucosylceramide-fraction was digested with recombinant lysosomal glucosidase to reduce β-glucosylceramide. The remaining material detected at the monohexosylceramide retention time was purified by preparative TLC and further resolved to upper and lower fractions. (B) TLC fractions were tested for activity by co-culture with the DN32 iNKT cell hybridoma and CD1d-transfected RAW-264.7 cells. (C) CD1d monomers loaded with the indicated lipids were assembled into tetramers that were used to stain BW58 cells transfected with an iNKT cell TCR or with a diverse NKT cell TCR that does not recognize α-galactosylceramide. Numbers indicate geometric mean fluorescence values. Data are representative of two independent experiments performed with separate starting material.

Fig. 2. TCR trap assay. (A) Synthetic α-glucosylceramide was loaded in recombinant CD1d molecules, mixed with recombinant iNKT cell TCRs (2C12), and run over a size-exclusion chromatography column. (B) Lipids were eluted from the early fraction excluded from the column and analyzed by linear ion trap MS. A dominant ion corresponding to the α-glucosylceramide standard was recovered and few other ions were detected. (C) The activity-enriched milk monohexosylceramide fraction was loaded in CD1d and treated as in (Fig. 2A). (D) The early fraction, CD1d alone, or TCR alone, were resolved on an SDS-PAGE gel and total proteins stained, yielding bands corresponding to TCR, CD1d, and β2M. Data are representative of two independent experiments performed with separate starting material.

Fig. 3. Linear ion trap MS analysis of CD1-lipid-TCR complexes. (A) Lipids extracted from the TCR trap assay fractions indicated Fig. 2C were analyzed by linear ion trap MS. Prominent ions from the late fraction were identified as (B) phosphatidylcholine, (C) sphingomyelin, (D) phosphatidylethanolamine, and (E) type 1 monohexosylceramide. (F) All major ions in the early fraction were an alkane series, and collision of 6 of the most abundant ions generated type 2 monohexosylceramide patterns. (G) Lipids were extracted from the early fraction, quantified, and co-cultured with the DN32 iNKT cell hybridoma and CD1d-transfected RAW-264.7 cells. Experiments were performed on lipids extracted two independent TCR trap assays.

Fig. 4. HPLC-MS analysis with Q-Tof MS detection. Lipids extracted from TCR trap assay fractions were analyzed by Q-ToF HPLC-MS experiments to provide high mass accuracy confirmation of the structures deduced in Fig. 3, measure lipid binding to TCR and to measure retention time in comparison with standards. HPLC tracings are shown for selected targeted ions based on predicted masses for the structures identified in Fig. 3. Experiments were performed on lipids extracted two independent TCR trap assays.

Fig. 5. Linear ion trap MS distinguishes α- from β-hexosylceramides. (A) Lipid standards were synthesized based on the structures of the type 2 monohexosylceramides and tested in co-culture with the DN32 iNKT cell hybridoma and CD1d-transfected RAW-264.7 cells. (B) MS-MS of sodiated adducts demonstrated characteristic fragmentation patterns for α- and β-glucosylceramide synthetic standards when studied over a range of collision energy settings. (C) Linear ion trap MS-MS was performed on the cow's milk unenriched, activity-enriched, and TCR trap early fraction lipids. Three experiments were performed for (A) and (B). Experiments in (C) were performed on lipids extracted two independent TCR trap assays.

Supporting Information

Supplementary figure legends

Fig. S1. One-dimensional proton NMR. Lipids eluted from the TCR trap early fraction were analyzed by one-dimensional proton NMR. A resonance with a chemical shift of 4.85 ppm was observed, characteristic of α-monohexosylceramide. This analysis was performed on one of two independent early fraction lipid eluates.

Fig. S2. HPLC-Q-ToF-MS collision spectra for type 1 and type 2 hexosylceramides. TCR trap assay (A) late fraction type 1 monohexosylceramides and (B) early fraction type 2 monohexosylceramides were analyzed by HPLC coupled to Q-ToF MS shown as extracted ion chromatograms (left) and CID-MS (right). Experiments were performed on lipids extracted two independent TCR trap assays and representative HPLC tracings are shown.

Fig. S3. Type 1 and 2 hexosylceramide ions determined by Q-tof MS. TCR trap assay late fraction type 1 (top panel) monohexosylceramides and early fraction type 2 (bottom panel) monohexosylceramides were analyzed by high mass accuracy Q-ToF MS. HPLC elution times of 3.4-5 min are shown, corresponding to the elution time of monohexosylceramide standards. Experiments were performed on lipids extracted two independent TCR trap assays, and representative HPLC tracings are shown.

Fig. S4. TCR trap early fraction co-elutes with a panel of synthetic monohexosylceramides. HPLC-Q-ToF-MS tracings of the synthetic monohexosylceramide standards with [M+H]+, *m/z* 786.682 (Fig. 5A) and TCR trap assay early fraction are shown. Experiments with synthetic lipid standards were performed at least three times. Experiments were performed on lipids extracted two independent TCR trap assays and representative HPLC tracings are shown.

Fig. S5. Linear ion trap MS distinguishes α- from β-galactosylceramides. As in Fig. 5B, α- and β-galactosylceramide synthetic standards were performed over a range of collision energies. Chemical

structures of α- and β-galactosylceramide with deduced fragmentations are shown. MS-MS of sodiated adducts demonstrated characteristic fragmentation patterns, like those observed for α- glucosylceramide and β-glucosylceramide synthetic standards. Three independent experiments were performed.

Fig. S6. Linear ion trap MS distinguishes α- from β-glycosylceramides with a sphingosine base. Using linear ion trap CID-MS, the synthetic standards with a d18:1 sphingosine base (structures shown) were analyzed. MS-MS tracings are shown. Three independent experiments were performed.

Fig. S7. Monohexosylceramides from calf thymus. (A) The monohexosylceramide fraction from calf thymus was purified by normal-phase TLC and analyzed by linear ion trap MS. Monohexosylceramides were detected in the thymus with the same *m/z* as those identified in cow's milk, including the type 2 monohexosylceramides *m/z* 724.7 and 808.8. (B) Calf thymus hexosylceramides were subjected to activity enrichment as in Fig. 1A,B. Collision of the m/z 724.7 and 808.8 before (top panels) and after activity enrichment (bottom panels) demonstrated an increase in loss of 162 u following activity enrichment, as observed for cow's milk (Fig. 5C). MS experiments were performed twice.

SI Methods and materials

Lipids

Starting with 50 mg of cow's milk glucocerebrosides (Matreya), two independent experiments were done to enrich the lipids in the starting material that activated iNKT cells (referred to as the "activity-enriched" fraction) (15). Briefly, lipids were first digested with recombinant human lysosomal glucocerebrosidase (Velaglucerase Alfa, Shire). The monohexosylceramides remaining after digestion were recovered by preparative TLC. The active lipid fraction was further isolated using the HICMW system (15). Lipids were resolved with normal phase TLC and visualized by staining with an α-napthol sulfuric acid (Sigma) solution and charring. Calf thymus (Research 87) monohexosylceramides were purified as previously described (15) and activity enrichment was carried out as for cow's milk lipids. β-glucosylceramide (d18:1/24:1) was purchased from Avanti polar lipids. Synthetic lipids, including α-galactosylceramide KRN7000 (d18:0-C26:0, containing a phytosphingosine base with a 26 carbon unsaturated *N*-acyl chain) were synthesized, α-glucosylceramide (d18:0-C22:0, containing a dihydrosphingosine base), β-glucosylceramide (d18:0-C22:0, containing a dihydrosphingosine base), α-galactosylceramide (d18:0-C22:0, containing a dihydrosphingosine base), and α-galactosylceramide (d18:0-C22:0, containing a dihydrosphingosine base), were synthesized as previously described (34).

CD1d tetramer staining of invariant and Type-2 NKT cell lines

Invariant NKT (Vα14Jα18 Vβ8.2Jβ2.1, CDR3α DRGSALGRLHFGAG, CDR3β GDAGGNYAEQFFGPG) and diverse NKT (Vα1Jα26 Vβ16Jβ2.1, CDR3α SEQNNYAQGLTFGLG, CDR3β SFWGAYAEQFFGPG) BW58 cell lines were generated using retroviral transduction (35). Recombinant mouse CD1d tetramers (36) were loaded with lipids dissolved in 0.5% Tyloxapol (Sigma) with Tris buffered saline pH 8.0 at a 3:1 lipid: protein molar ratio. Cell lines with invariant and diverse NKT cell TCRs co-stained with TCRβ (clone H57-597) and CD1d tetramers loaded with C24:1 α-galactosylceramide (PBS44) (kindly provided by Paul Savage, Brigham Young University), α- glucosylceramide, activity enriched milk antigen or unloaded CD1d tetramer.

TCR trap assay

Recombinant Vα14Jα18 Vβ8.2Jβ2.5 TCR (2C12) TCR, CDR3β GDEGYTQYFGPG was prepared as previously described for the production of recombinant Vα14Jα18 Vβ7 TCR (37). Following lipid loading, excess Tyloxapol was removed from recombinant mouse CD1d using ion exchange chromatography (Mono Q 10/100GL, GE Healthcare). Mouse CD1d and soluble 2C12 TCR were passaged separately over a Superdex-200 16/60 gel-filtration column (GE healthcare), just prior to the generation of ternary complexes, to remove any protein aggregates. CD1d-milk Ag-TCR and CD1d-α-glucosylceramide-TCR ternary complexes were generated by mixing equal amounts of CD1d and by passage over Superdex-200 column at a flow rate of 0.5mls/min. Gel filtration fractions were analyzed on a 12% SDS-PAGE gels (Biorad) using Coomassie blue stain.

iNKT cell activation assay

Lipid activity assays were performed with the DN32 iNKT cell hybridoma (38) and CD1d-transfected RAW-264.7 cells (39). Briefly, lipids were dried under a nitrogen stream, then sonicated in serum-containing assay media. 5 $x10^4$ DN32 cells and 5 $x10^4$ CD1d-transfected RAW-264.7 cells were co-cultured with lipids for 16 hr at 37 ºC. Assays were performed in RPMI-1640 media supplemented with 10% fetal calf serum (Gemini Bio), penicillin/streptomycin, L-glutamine and β-mercaptoethanol (Life Technologies) at 37ºC in 5% CO_2. For X-Y graphs, standard error of the mean is shown.

Nanospray ESI-MS

All organic solvents were HPLC grade. Lipids were extracted from CD1d proteins were with chloroform, methanol and water according to the method of Bligh and Dyer (40). Briefly, protein samples in 200 µl 10mM Tris pH 8.0 150 mM NaCl were transferred to 15 ml glass tubes, treated with 0.75 ml of chloroform:methanol (1:2 (v:v)), vortexed for 1 min to denature the protein and extract lipids. The mixture was treated with 0.25 ml of chloroform and vortexed for 30 seconds, and then treated with 0.25 ml of water and vortexed for another 30 seconds. The mixture was centrifuged at 850 x g for 10 min to separate the layers. The upper aqueous layer

was removed, and the lower organic lipid containing layer was transferred to another clean glass tube. The organic solvents were dried under a liquid nitrogen stream and the lipid residue was re-dissolved in chloroform/methanol (1:2) and stored at -20 °C. For a shotgun lipid analysis, the extracted lipids were loaded onto a nanospray tip for electrospray mass spectrometry (ESI-MS) and multistage collision-induced dissociation tandem mass (CID-MS) using linear ion-trap mass spectrometer (LXQ, Thermo Scientific). Collision energy was 20–30% of maximum and product ions were trapped with a *q* value of 0.25.

HPLC-MS analysis

To obtain the accurate mass and retention time information, the extracted lipids from 25 – 400 μg of CD1d protein and normalized to 25 μM based on the concentrations of input protein. The normalized samples (25 μl) were dried and re-dissolved in 25 μl of starting mobile phase B (see below), and 20 μl were injected for HPLC-ESI-MS analysis (Agilent 6520 Accurate-Mass Q-TOF and 1200 series HPLC system using a normal phase Varian MonoChrom Diol column, 3 μm x 150 mm x 2 mm, and running at 0.15 ml/min) according to the published methods (41). The solvent mobile phases were (A) isopropanol/methanol (70:30, (v:v)) and (B) hexane/isopropanol (70:30, (v:v)). Both mobile phases were supplemented with 0.1% formic acid and 0.05% ammonium hydroxide. The gradient system was following: 0-10 min, 100% B; 10-17 min, 100% to 50% B; 17-22 min, 50% B; 22-30 min, 50% to 0% B; 30-35 min 0% B; 35-40 min, 0% to 100% B; 40-44 min, 100% B; followed by additional 6 min 100% B post-run. The glycolipid concentrations in the extracted fractions were estimated by comparing the peak areas of the extracted ion chromatograms to those of the external standard curves (β-glucosylceramide d18:1/24:1). All monohexosylceramide variants were assumed to have similar response factor for ionization. For the synthetic glucosyl- or galactosylceramides, the normal phase HPLC elution system was optimized for resolution of individual compounds using an isocratic condition with mobile phase solvents of hexane/isopropyl alcohol (70:30 (v:v)), containing formic acid 0.1% (v:v) and ammonium hydroxide 0.05% (v:v). The run was set at 0.15 ml/min for 10 min. The synthetic compounds were eluted from 4 to 6 min. CID-MS experiments were performed with collision energy of 30 V and 1.3 *m/z* isolation width.

NMR analysis

NMR was performed at the University of Georgia Complex Carbohydrate Research Center as previously described (15). Samples were dissolved in 2:1 CDCl3-CD3OD followed by one-dimensional proton NMR acquired on a Varian VNMRS 900 MHz spectrometer at 27ºC.

Fig. 1

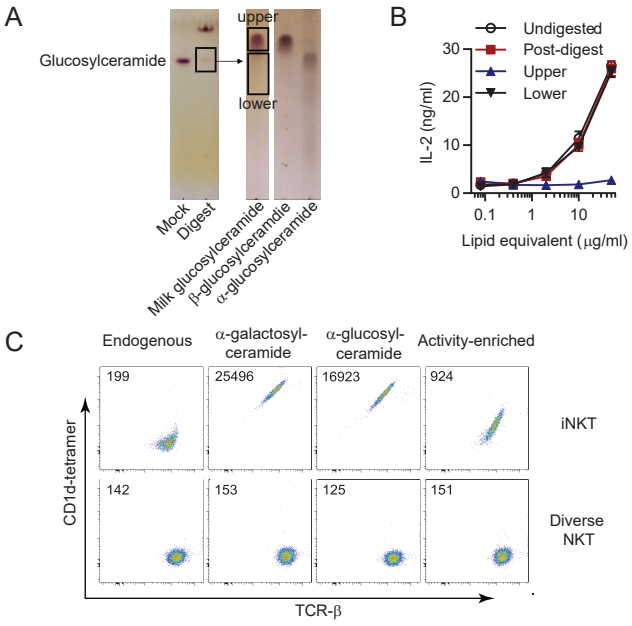

A
Glucosylceramide
upper
lower
Mock
Digest
Milk glucosylceramide
β-glucosylceramide
α-glucosylceramide
B
Undigested
Post-digest
Upper
Lower
IL-2 (ng/ml)
30
20
10
0
0.1
1
10
Lipid equivalent (μg/ml)
C
Endogenous
α-galactosyl-ceramide
α-glucosyl-ceramide
Activity-enriched
199
25496
16923
924
142
153
125
151
iNKT
Diverse NKT
CD1d-tetramer
TCR-β

Fig. 2

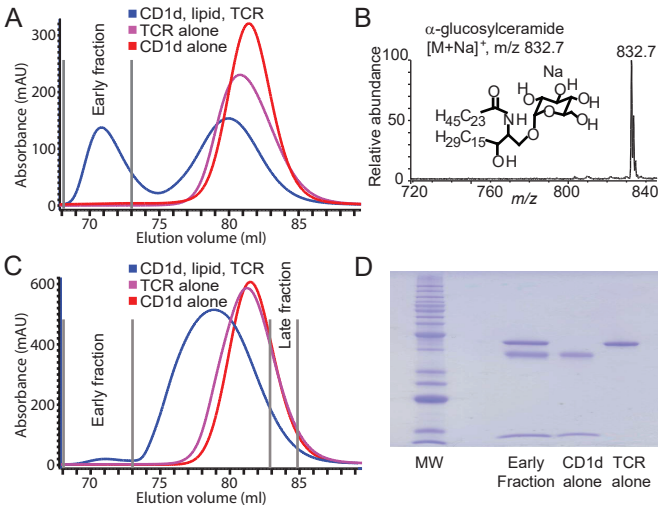

A
CD1d, lipid, TCR
TCR alone
CD1d alone
Early fraction
Absorbance (mAU)
Elution volume (ml)
B
α-glucosylceramide
[M+Na]+, m/z 832.7
832.7
Relative abundance
m/z
C
CD1d, lipid, TCR
TCR alone
CD1d alone
Early fraction
Late fraction
Absorbance (mAU)
Elution volume (ml)
D
MW
Early Fraction
CD1d alone
TCR alone

Fig. 3

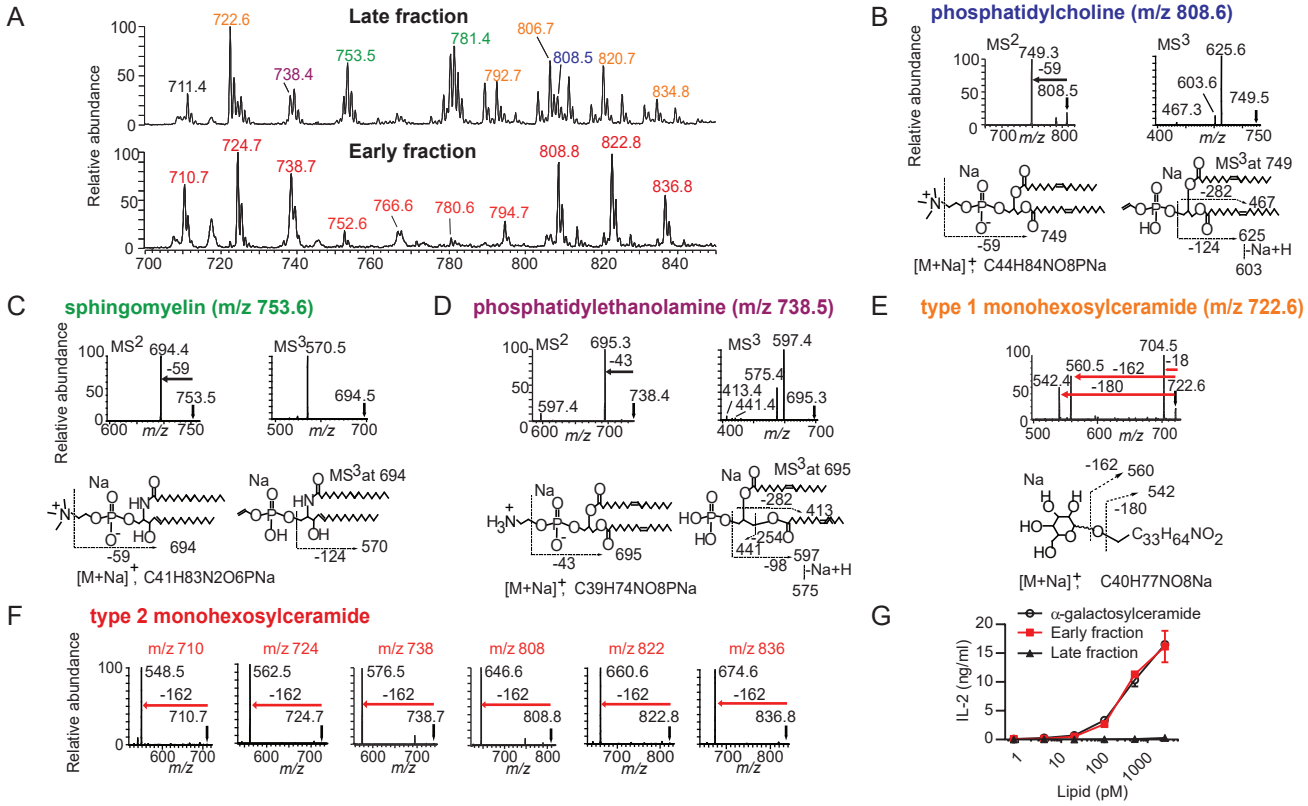

A
Late fraction
Early fraction
Relative abundance
B
phosphatidylcholine (m/z 808.6)
C
sphingomyelin (m/z 753.6)
D
phosphatidylethanolamine (m/z 738.5)
E
type 1 monohexosylceramide (m/z 722.6)
F
type 2 monohexosylceramide
m/z 710
m/z 724
m/z 738
m/z 808
m/z 822
m/z 836
G
α-galactosylceramide
Early fraction
Late fraction
IL-2 (ng/ml)
Lipid (pM)

Fig. 4

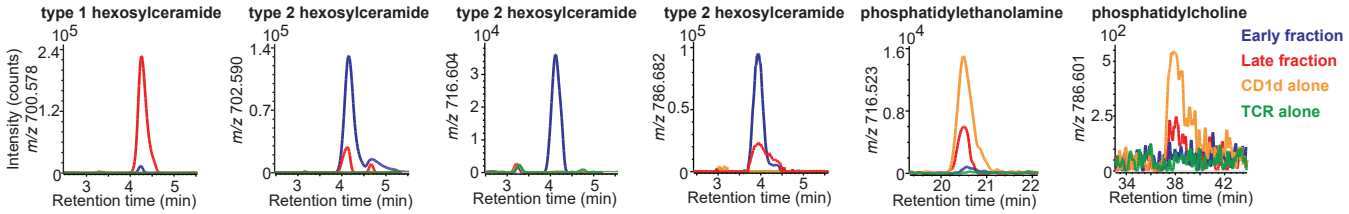

type 1 hexosylceramide
type 2 hexosylceramide
type 2 hexosylceramide
type 2 hexosylceramide
phosphatidylethanolamine
phosphatidylcholine
Intensity (counts)
m/z 700.578
m/z 702.590
m/z 716.604
m/z 786.682
m/z 716.523
m/z 786.601
Retention time (min)
Early fraction
Late fraction
CD1d alone
TCR alone

Fig. 5

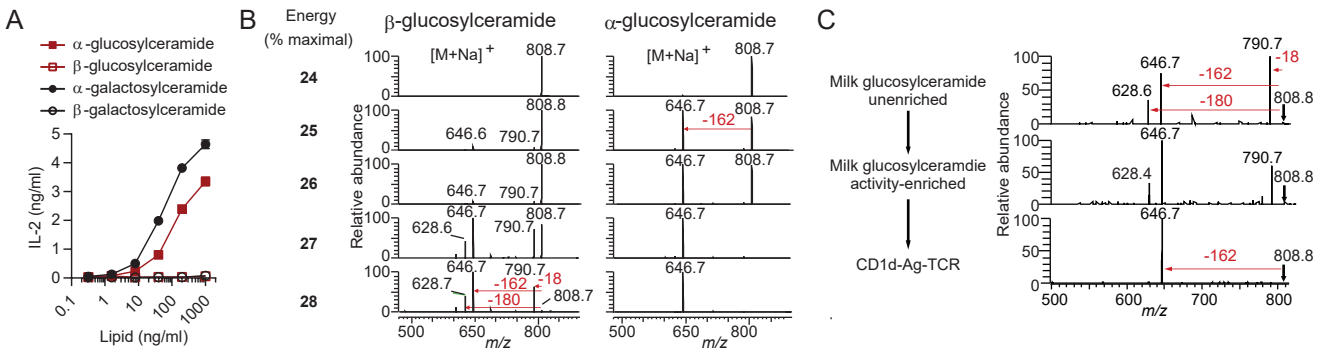

A
α-glucosylceramide
β-glucosylceramide
α-galactosylceramide
β-galactosylceramide
IL-2 (ng/ml)
Lipid (ng/ml)
B
Energy
(% maximal)
β-glucosylceramide
α-glucosylceramide
[M+Na]+
Relative abundance
m/z
C
Milk glucosylceramide
unenriched
Milk glucosylceramdie
activity-enriched
CD1d-Ag-TCR

Fig. S1

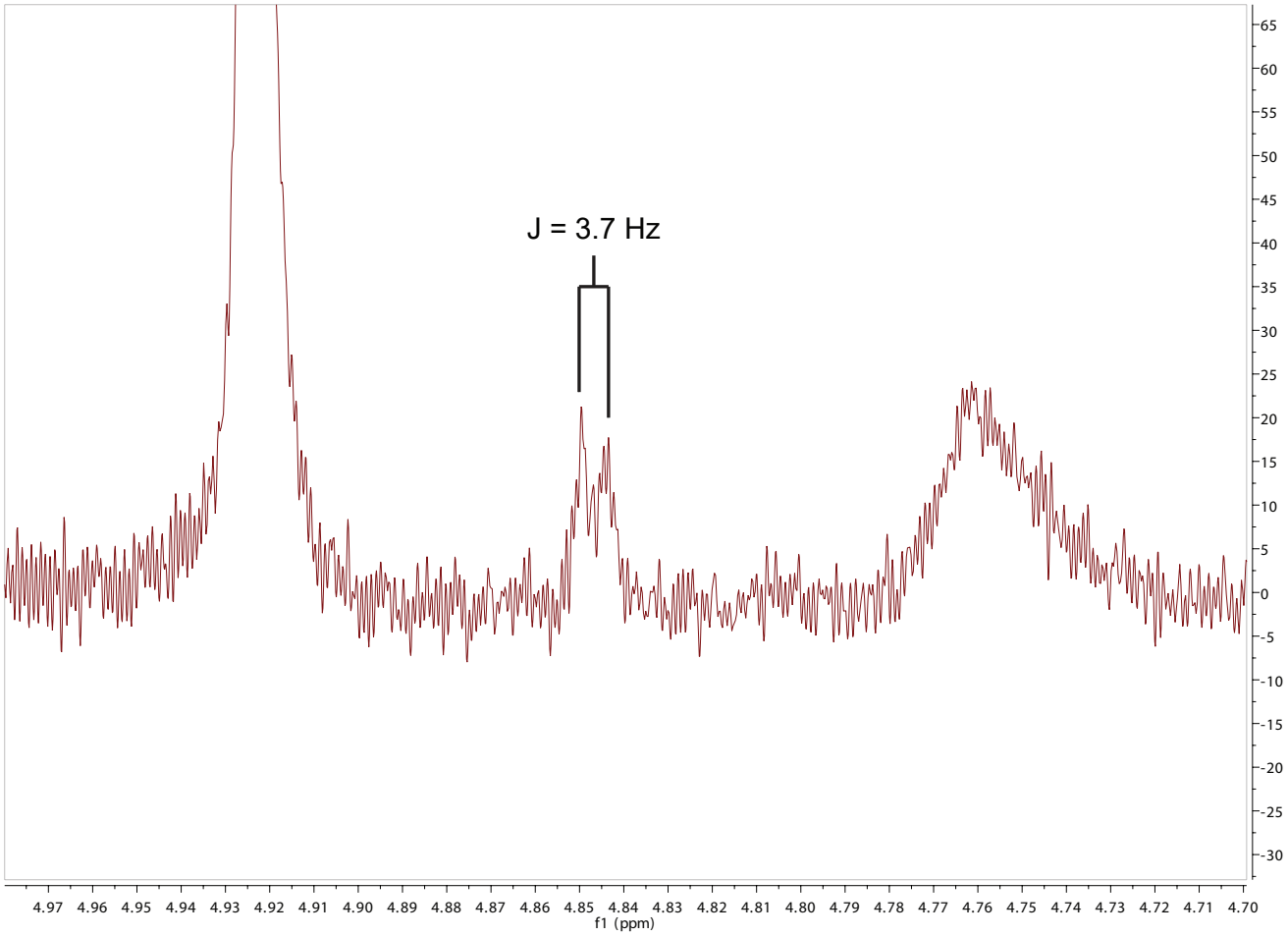

J = 3.7 Hz
65
60
55
50
45
40
35
30
25
20
15
10
5
0
-5
-10
-15
-20
-25
-30
4.97 4.96 4.95 4.94 4.93 4.92 4.91 4.90 4.89 4.88 4.87 4.86 4.85 4.84 4.83 4.82 4.81 4.80 4.79 4.78 4.77 4.76 4.75 4.74 4.73 4.72 4.71 4.70
f1 (ppm)

Fig. S2

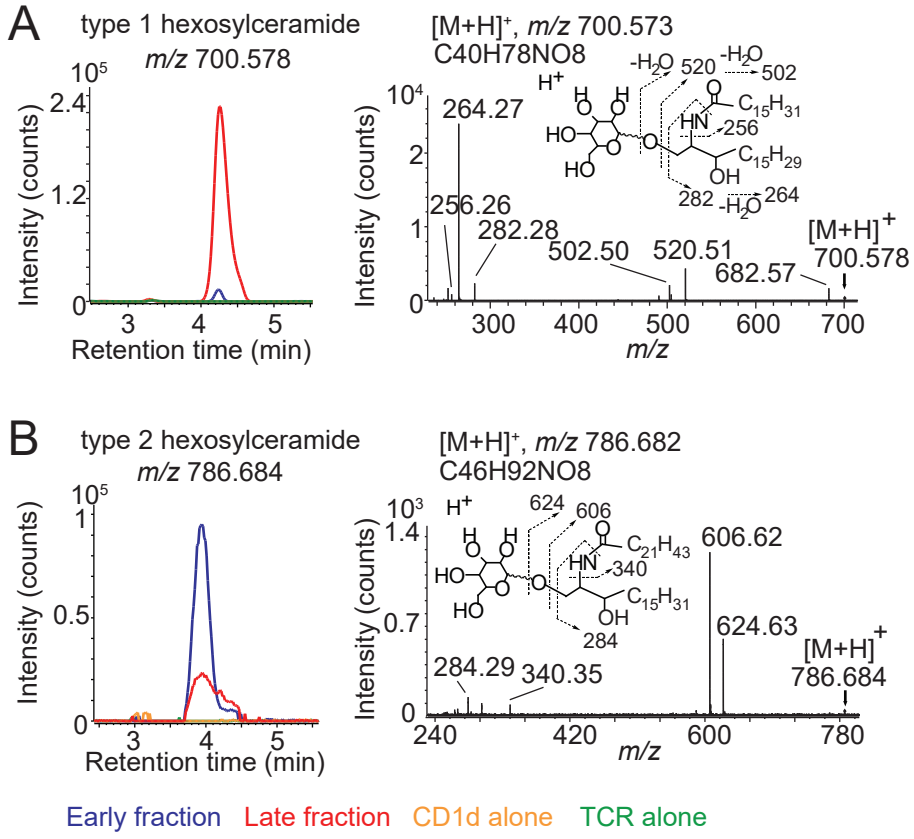

A
type 1 hexosylceramide
m/z 700.578
[M+H]+, m/z 700.573
C40H78NO8
Intensity (counts)
Retention time (min)
m/z
264.27
256.26
282.28
502.50
520.51
682.57
[M+H]+
700.578
B
type 2 hexosylceramide
m/z 786.684
[M+H]+, m/z 786.682
C46H92NO8
606.62
624.63
284.29
340.35
[M+H]+
786.684
Early fraction
Late fraction
CD1d alone
TCR alone

Fig. S3

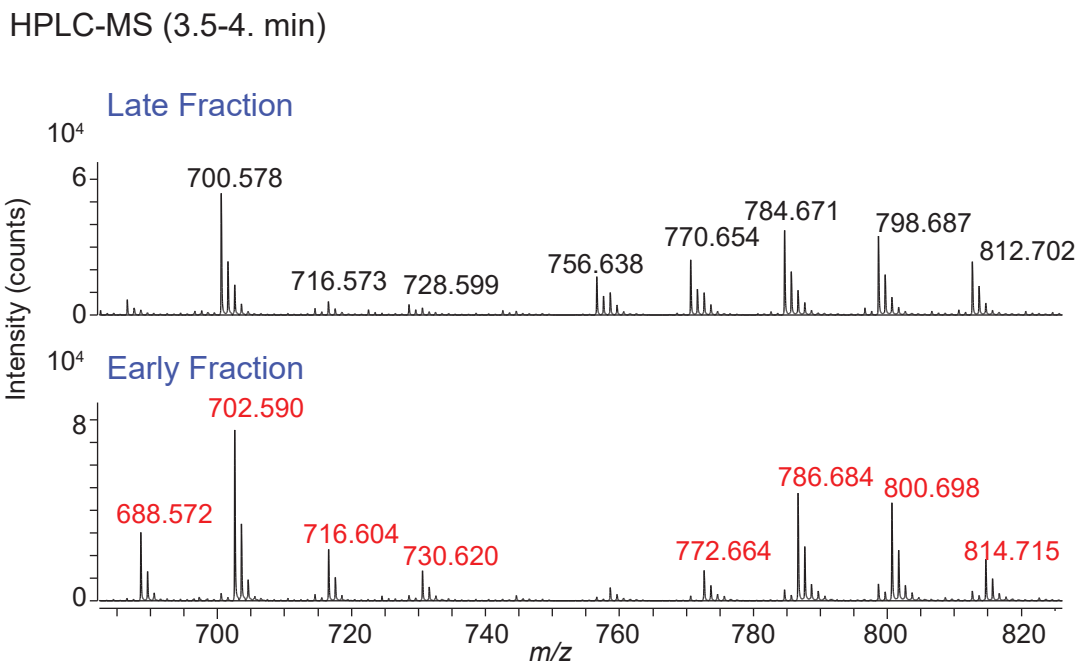

HPLC-MS (3.5-4. min)
Late Fraction
10⁴
6
0
700.578
716.573
728.599
756.638
770.654
784.671
798.687
812.702
Intensity (counts)
10⁴
Early Fraction
8
0
702.590
688.572
716.604
730.620
772.664
786.684
800.698
814.715
700
720
740
760
780
800
820
m/z

Fig. S4

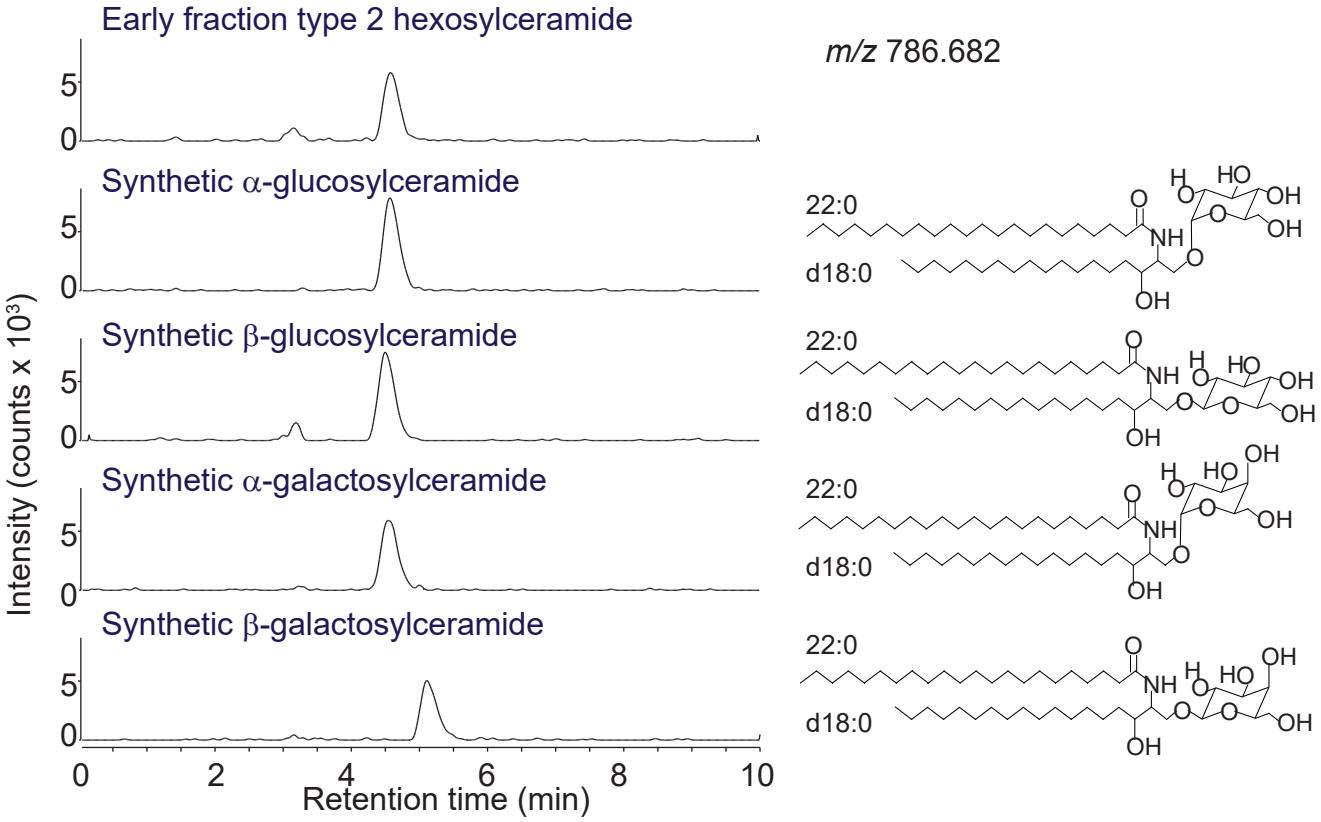

Early fraction type 2 hexosylceramide
Synthetic α-glucosylceramide
Synthetic β-glucosylceramide
Synthetic α-galactosylceramide
Synthetic β-galactosylceramide
Intensity (counts x 10³)
Retention time (min)
m/z 786.682
22:0
d18:0

Fig. S5

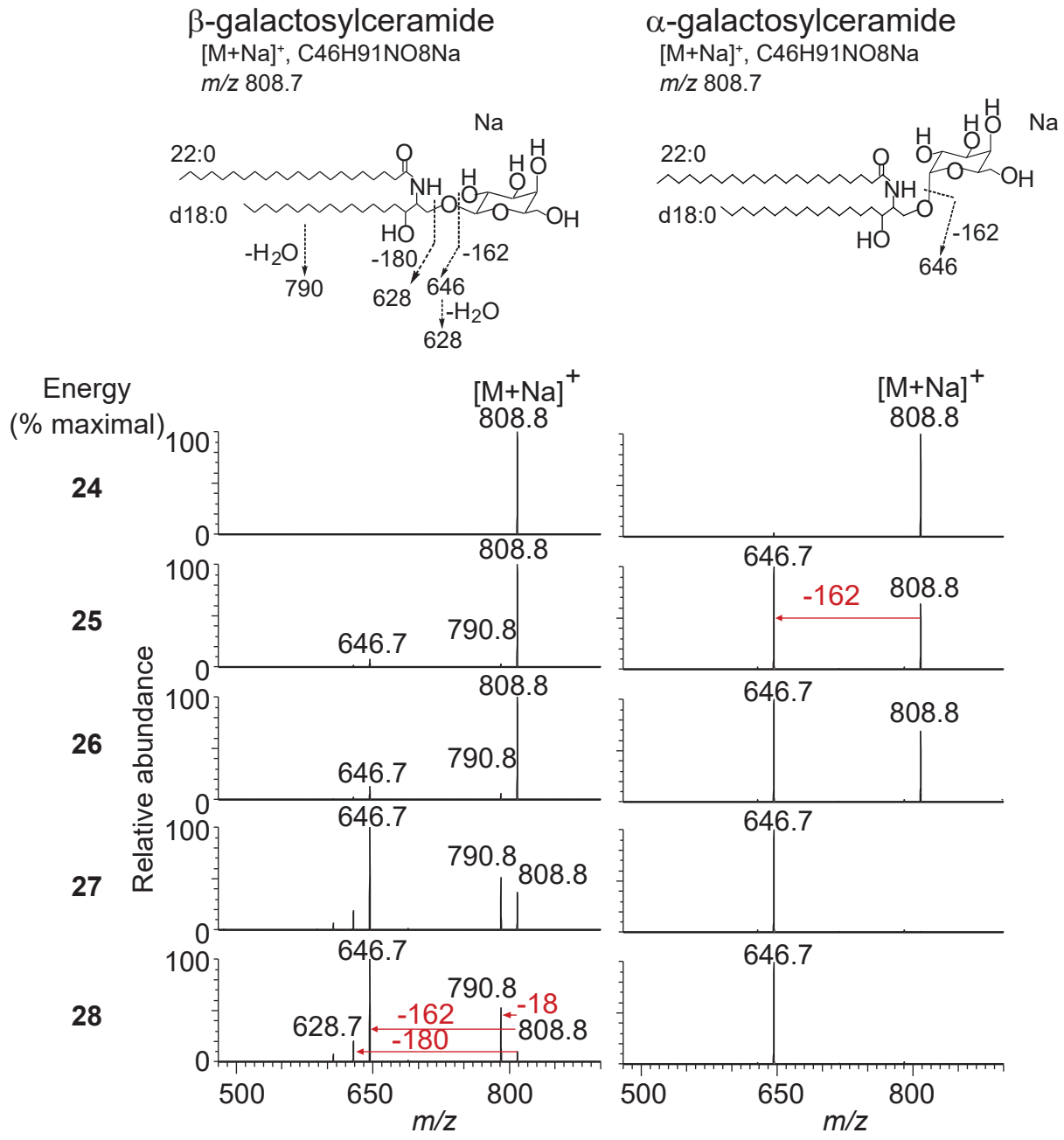

β-galactosylceramide
[M+Na]+, C46H91NO8Na
m/z 808.7
α-galactosylceramide
[M+Na]+, C46H91NO8Na
m/z 808.7
22:0
d18:0
Na
NH
HO
OH
-H2O
790
-180
628
-162
646
-H2O
628
Energy
(% maximal)
[M+Na]+
808.8
24
25
26
27
28
Relative abundance
100
0
646.7
790.8
-162
-18
628.7
-180
500
650
800
m/z

Fig. S6

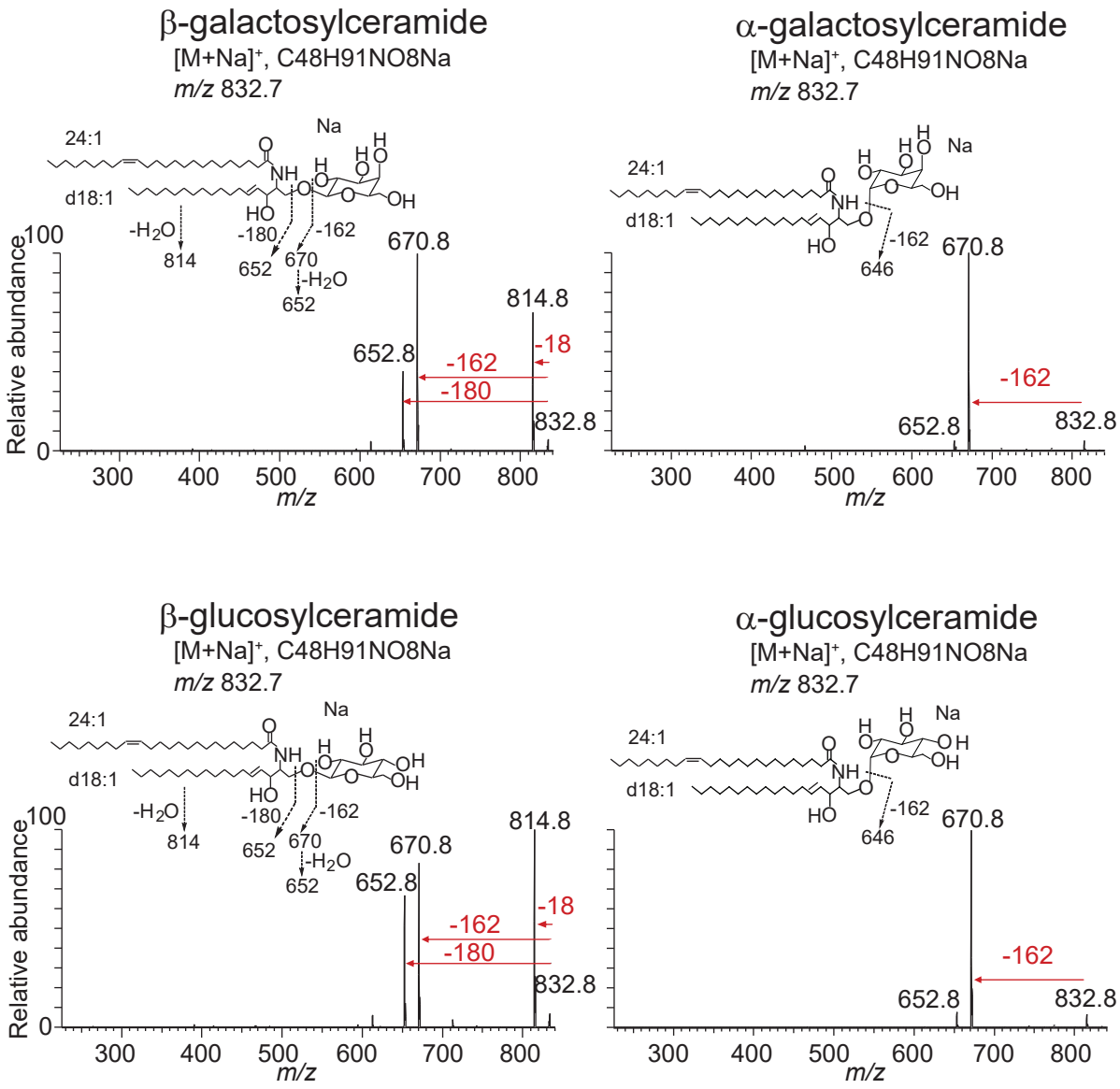
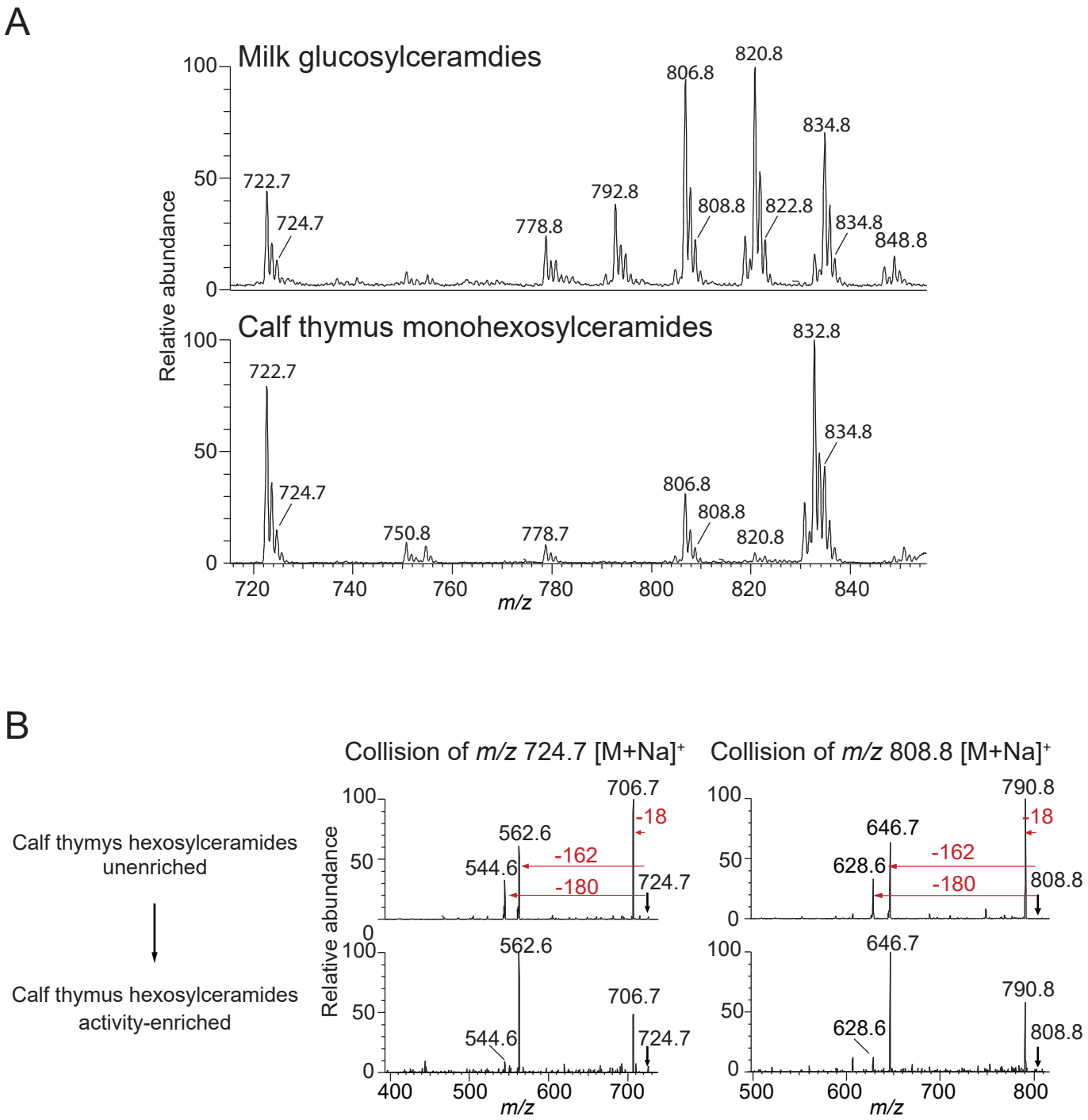

β-galactosylceramide
[M+Na]+, C48H91NO8Na
m/z 832.7
24:1
d18:1
Na
-H2O
814
-180
652
-162
670
-H2O
652
670.8
814.8
652.8
-18
-162
-180
832.8
Relative abundance
100
0
300
400
500
600
700
800
m/z
α-galactosylceramide
[M+Na]+, C48H91NO8Na
m/z 832.7
24:1
d18:1
Na
-162
646
670.8
652.8
-162
832.8
m/z
β-glucosylceramide
[M+Na]+, C48H91NO8Na
m/z 832.7
24:1
d18:1
Na
-H2O
814
-180
652
-162
670
-H2O
652
670.8
814.8
652.8
-18
-162
-180
832.8
Relative abundance
100
0
300
400
500
600
700
800
m/z
α-glucosylceramide
[M+Na]+, C48H91NO8Na
m/z 832.7
24:1
d18:1
Na
-162
646
670.8
652.8
-162
832.8
m/z

Fig. S7


A
Milk glucosylceramdies
Calf thymus monohexosylceramides
Relative abundance
m/z
722.7
724.7
778.8
792.8
806.8
808.8
820.8
822.8
834.8
848.8
750.8
778.7
832.8
B
Calf thymys hexosylceramides unenriched
Calf thymus hexosylceramides activity-enriched
Collision of m/z 724.7 [M+Na]+
Collision of m/z 808.8 [M+Na]+
706.7
-18
562.6
-162
544.6
-180
724.7
790.8
646.7
628.6
808.8